Exhibit A

FUNDAMENTAL VIROLOGY

Third Edition

Editors-in-Chief

Bernard N. Fields, M.D.
Department of Microbiology and Molecular Genetics
Harvard Medical School
Boston, Massachusetts

David M. Knipe, Ph.D.
Department of Microbiology and Molecular Genetics
Harvard Medical School
Boston, Massachusetts

Peter M. Howley, M.D.
Department of Pathology
Harvard Medical School
Boston, Massachusetts

Associate Editors

Robert M. Chanock, M.D.
Laboratory of Infectious Diseases
National Institute of Allergy and Infectious Disease
National Institutes of Health
Bethesda, Maryland

Joseph L. Melnick, Ph.D., Sc.D.
Department of Virology and Epidemiology
Baylor College of Medicine
Texas Medical Center
Houston, Texas

Thomas P. Monath, M.D.
Ora Vax, Inc.
Cambridge, Massachusetts

Bernard Roizman, Sc.D.
Department of Molecular Genetics and Cell Biology
University of Chicago
Chicago, Illinois

Stephen E. Straus, M.D.
Laboratory of Clinical Investigation
National Institutes of Health
Bethesda, Maryland

Lippincott - Raven
PUBLISHERS
Philadelphia • New York

Fundamental Virology, Third Edition
edited by B.N. Fields, D.M. Knipe, P.M. Howley, et al.
Lippincott - Raven Publishers, Philadelphia © 1996

CHAPTER 15

Bacteriophages

Allan M. Campbell

HISTORY

Bacterial viruses (generally called bacteriophages or simply phages) were discovered independently by Twort (66) and d'Herelle (33) as filtrable, transmissible agents of bacterial lysis. Although their relevance to the study of animal and plant viruses was not universally accepted, phages proved so tractable to experimentation that information about their nature was soon forthcoming. For example, the ability of phage to form isolated plaques on bacterial lawns was noted; and d'Herelle correctly attributed the linearity of plaque assays to the self-replicating nature of the phage particles (as expected for a subcellular pathogen) rather than to the cooperative effect of many particles (as expected for a typical poison.)

The early hope that phage might be useful in combating pathogenic bacteria (either therapeutically or in decontamination of the human environment, such as in purifying water supplies) did not materialize, although occasional therapeutic applications are made. They are most effective where susceptible bacteria are concentrated in lesions or abscesses, where large amounts of phage can be injected. Starting in the 1920s, a more successful and widespread use of phage was in the typing of bacterial strains of medical interest by testing their sensitivity to a battery of phages.

A new era in phage research was initiated in the early 1940s by Max Delbrück, Salvador Luria, and their disciples. Their efforts laid the groundwork for the use of phage in the fundamental experiments of molecular biology; but the work itself is notable for its deemphasis of biochemistry and its focus on basic biological information relevant to the mechanism of self-replication. The techniques used were simple, mainly microbiological, and some of the questions investigated had already received indicative answers in the previous two decades. What was new was an emphasis on sharp, incisive answers, quantitative where appropriate, and the development of a few standardized systems suitable for obtaining such answers. Conditions were determined for the rapid, reproducible attachment of phage to bacterial cells, allowing synchronous infection of all the cells in a culture with the resulting "one-step" growth. The effect of varying the ratio of phage to bacteria (multiplic-

A.M. Campbell: Department of Biological Sciences, Stanford University, Stanford, CA 94305

ity of infection) was analyzed as a problem in random statistics, to which the Poisson distribution could properly be applied. It was also shown that synchronously infected cells pass through an eclipse phase, where infectious phage is unrecoverable from infected cells, and that infectious particles appear at a later time (2). The ability of one phage to exclude another and of coinfecting phage to interact productively either genetically (by recombination) or phenotypically (by complementation or phenotypic mixing) or both (as in the production of viable phage from cells multiply infected with radiation-damaged particles) was rigorously documented. Quantitative analysis at a comparable level was introduced into animal virology with the development of plaque assays on tissue monolayers (21). Thus, the foundations of modern animal virology (independent of its molecular aspects) go back to the phage work of the 1940s.

Studies of phage genetics started in the 1940s with the construction of recombinational maps (36) and continued in the 1950s with Benzer's analysis of fine structure genetics of the T4*r*II genes (8) and with the discovery and use of conditionally lethal mutations in λ and T4, which allowed dissection of the developmental programs of phages (12,23). The *r*II work depended conceptually on the prior discovery of phenotypic mixing (53), which showed that when mutants of the same phage coinfect a cell, phage gene products are drawn from a common pool, and the conditional lethal work depended on the operational protocol laid out in the *r*II studies.

Serious phage biochemistry was initiated in the 1930s, when Schlesinger determined that phage particles were about 50% DNA and 50% protein (60). It did not enter into the mainstream of phage work until the 1950s, when Hershey and Chase showed by radiolabeling that the phosphorus (DNA) entered the cell during T2 infection but that most of the sulfur (protein) did not (35). The experiment followed on the demonstrations of the eclipse phase and electron microscopy, showing that empty phage heads remain attached to the cell envelope and that the DNA of virions could be separated from the protein shell by osmotic shock (1). Concurrently, the enzymology of T4 infection was developed in detail, clearly indicating that the phage was introducing instructions for proteins present neither in the uninfected cell nor in the virion (17).

It was reported in the 1920s that many bacterial strains harbored phage permanently and secreted it into the medium, so that every culture contained phage. Such strains are called lysogenic. The Luria-Delbrück school initially chose to ignore lysogeny, in part because of the possibility that the phenomenon might be explainable simply by the adherence of extracellular phage particles to the bacterial surface and their occasional detachment from it. Increased rigor was introduced by Lwoff et al. around 1950 in a series of experiments showing first, that when individual cells were allowed to divide and daughter cells separated into different drops of liquid by micromanipulation, both daughters retained the ability to produce phage; second, that an occasional cell lysed and liberated a burst of phage (so that it was unnecessary to postulate secretion by living cells); and third, that certain treatments (such as ultraviolet light or other agents that induce the SOS response) induce mass lysis and phage liberation by all cells of the culture (49). As with infected cells in the eclipse phase, no infectious phages were found when lysogenic cells were artificially disrupted.

Phages capable of producing lysogenic complexes are called temperate, and the noninfectious form in which phage genetic information is transmitted is called prophage. Of the temperate phages, coliphage λ has been studied most intensively, partly because it grows in the K-12 strain of *Escherichia coli*, where conjugational and transductional mapping were possible by the 1950s. It was soon shown that prophage has a specific location on the bacterial map, and later that its genome is linearly inserted into the continuity of the bacterial chromosome (13,43).

From 1950 onward, the virtues of a simple identifiable genome with rigorous genetics that could be simultaneously introduced into (or induced to replicate in) all the cells of a culture made phage the favored experimental material for many of the classical experiments of molecular biology, such as the demonstration of DNA breakage and joining during genetic recombination and the proof of the triplet code. In the 1960s, λ, in particular, become the favored object for study of the genetic hierarchy controlling a simple developmental program. Although the initial investigations on conditionally lethal mutations were pursued in far more depth with T4 than with λ, λ had certain inherent advantages: the initiation of the entire program of the lytic cycle from two promoters directly controlled by repressor, the uniformity of the DNA molecules extracted from virions, and the specific probes for different parts of the genome provided by specialized transducing phages and deletions or substitutions (whose locations could be seen via heteroduplex mapping) made the effects of genetic manipulation on molecular events readily interpretable; and the existence of alternative pathways toward lysis or lysogeny added to the interest.

Since 1970, phage work merges with the rest of biology because the invention of artificial cloning allows the application to many systems of the techniques early available for phage. The background data on phage systems, plus their facile manipulation, continues to allow their study at a high level of sophistication.

VIRULENT PHAGES

Like viruses in general, phages can be divided into those with RNA genomes (mostly small and single stranded), those with small DNA genomes (generally less than 10 kb, mostly single stranded), and those with medium to large DNA genomes (30–200 kb). The last group includes most

of the temperate phages, as well as the first virulent phages to be studied intensively.

Large DNA Phages

The classical T coliphages selected for study by the Luria-Delbrück school fall into four groups: T1 (about 50 kb); T2, T4, and T6 (about 170 kb); T3 and T7 (40 kb); and T5 (about 130 kb). The members of a group have similar virion morphology and can produce viable recombinants in mixed infection.

Phage T4

Virion

T4, the object of many classical experiments, is an appropriate prototype. More than 200 T4 genes have been identified through mutational studies. The virion contains 43 phage-encoded proteins: 16 are located in the head that encapsidates the DNA and 27 form the tail, through which DNA passes into the cell during infection. The head is an elongated T = 13 icosahedron with an extra equatorial row of capsomeres. The tail consists of a hollow core (surrounded by a contractile sheath) that terminates in a baseplate to which are attached six fibers. The whole apparatus functions as a syringe for injection of phage DNA into the interior of the cell. The tips of the tail fibers make initial contact with the lipopolysaccharide surface receptors on the bacterium. Once the phage is anchored by tail fiber attachment, random motion brings the baseplate in contact with the cell surface. Either this contact or tail fiber attachment itself triggers conformational changes in the tail: the center of the baseplate opens like a shutter, allowing the tip of the core to pass through, and the sheath contracts, driving the core through the cell envelope, after which DNA is released into the cell. The double-stranded DNA is distinctive in containing hydroxymethylcytosine (heavily glucosylated) rather than cytosine. It is cyclically permuted and terminally repetitious; i.e., the sequences of various molecules can be represented as ABCDEFGHAB, DEFGHABCDE, GHABCDEFGH, etc. This is most simply demonstrated by allowing the DNA to separate into single strands and then reannealing with complementary strands from other virions, generating circular duplexes with single-stranded tails.

DNA Transactions

Infection triggers a massive degradation of cytosine-containing host DNA with subsequent enzymatic conversion of deoxycytosine monophosphate to deoxyhydroxymethylcytosine monophosphate. The phage encodes a battery of enzymes that perform these and other conversions of deoxyribonucleotides that are then used in phage DNA replication. Bidirectional replication is initiated at several potential replication origins along the DNA, and elongation is performed by a complex of phage-encoded enzymes, leading and lagging strand synthesis proceeding coordinately from a single replication complex. Because there is no mechanism for priming synthesis at the extreme 5′ end of the lagging strand, this first round of replication produces linear molecules with protruding 3′ ends.

Soon after infection, these early replication origins cease to function (because replication there requires priming by RNA made by host RNA polymerase, whose promoter recognition specificity is altered). Late replication initiates at recombination intermediates formed by invasion of 3′ ends into homologous double-stranded DNA. Because of the terminal redundancy, such homologous sequences are available, even in single infection. As replication proceeds, repeated invasions of this kind generate a complex network of intracellular viral DNA, which includes end-to-end concatemers (the preferred packaging substrate for T4). In packaging, empty heads are first assembled, then filled with DNA. Cutting is coordinated with packaging, so that headful lengths are cut to size after the head is filled. The packaging length exceeds the genome length, hence the terminal redundancy. The position of molecular ends along the concatemer is close to random.

This picture of the DNA transactions has genetic consequences; in fact, much of it was inferred from genetic studies that preceded the biochemistry and molecular biology. First, the recombination rate in T4 infection is very high, producing a linkage map thousands of centimorgans in length. It was early noted that this corresponded to about one recombination per replication cycle. Second, the linkage map is circular, as expected if virion DNA has random endpoints. Third, when cells are infected with phages of two different genotypes, some of the progeny particles can carry different information within the terminal duplication; for example, in a mixed infection between T4 and T4*h*, some progeny plaques contain phage of both *h* and *h*$^+$ genotypes. Another mechanism that can also generate such mixed plaques is packaging of heteroduplex DNA recombination intermediates. Both types of heterozygous progeny are in fact demonstrable (63).

Finally, there are the genetic implications of headful packaging. If the length of the terminal overlap represents the difference between packaging length (determined by head size) and genome length, then that length should change in a predictable manner if either packaging length or genome length is deliberately altered. The first test was to alter genome length with a nonlethal deletion. The prediction that this would increase the length of the terminal overlap (and therefore the fraction of particles heterozygous for any locus in the genome) was fulfilled (65). Later experiments with other phages that use headful packaging, such as P22, showed that increasing the genome length by insertion of extra DNA has the reverse effect of decreasing the terminal overlap.

The converse experiment, where genome length remains constant but packaging length changes, was accomplished through examination of virions with abnormal morphology. Preparations of T4 contain a small fraction of virions whose heads are isometric rather than elongated, separable from the majority type by ultracentrifugation. Certain treatments (such as phage development in the presence of amino acid analogs) induce formation of giant particles with the same diameter as normal virions but additional rows of capsomeres, approaching helical tubes in shape. Missense mutations in the major capsid protein can increase the proportion of both isometric and giant particles (19). Both types of abnormal particle inject their DNA into infected cells. The packaging capacity of isometric particles is about 67% that of normal virions; accordingly, single infection is unproductive, but multiple infection allows a normal cycle through recombination between DNA molecules of less than genome length with random endpoints. Giant particles have longer than normal DNA (up to several genome lengths) and a correspondingly high degree of heterozygosity when coming from mixed infection. They are also highly resistant to ultraviolet light because of complementation and recombination between damaged genomes.

Regulation

The genes whose products are needed for phage DNA synthesis and host DNA breakdown, including those mediating nucleotide metabolism and the seven proteins that make up the replication complex, are all expressed immediately after infection, being transcribed from promoters recognizable by *E. coli* RNA polymerase with σ70. With time, early synthesis is shut off and other genes that code for virion components and lysis genes are activated. The shutoff of early genes is effected both at the transcriptional and the translational level. Translational control is sometimes autogenous, as for the DNA polymerase gp43 and the single-stranded binding protein gp32; many other early genes are controlled by a dedicated translational repressor, product of the *regA* gene, itself an early gene subject to autogenous control (3,11,71).

Transcriptional control operates on three different types of T4 promoters (28). The early promoters have sequences approximating the consensus for *E. coli* σ70 (with a small but apparently real difference in consensus). Middle promoters require *E. coli* polymerase with σ70, in addition to a T4-specified positive regulator, *motA*. Their hallmark is a sequence TGCTT at around –32. Late promoters follow –12 consensus TATAAATA. They are transcribed by *E. coli* holoenzyme, with a phage-coded σ factor (gp55) replacing σ70. In normal infection, late promoter activation requires concurrent DNA replication. *In vitro*, activation can be achieved by nicks or gaps on the nontranscribed strand near the promoter; treatments that can create damage *in vivo* can also induce some replication-independent transcription. Thus, the timing of late transcription is prescribed by the requirement for phage-specific factors (including gp55) and also by the modification of the template through introduction of interruptions during DNA synthesis.

In vivo, late transcription requires the integrity of all the DNA replication genes. *In vitro*, activation of late transcription by nicks or gaps requires three replication proteins gp45 and gp42/66, which are associated with the phage-coded polymerase gp43 during replication. An RNA polymerase-associated protein, gp33, is also required. An attractive model is that gp45 and gp42/66 recognize the DNA interruption and gp33 receives the enhancement signal thus created. Unlike other enhancers, transcription cannot be stimulated by bringing the DNA interruption spatially close to the promoter through catenation of DNA molecules; enhancement is only seen when all signals are on the same DNA double helix (34).

Some other ancillary changes that occur during infection may reinforce the temporal sequence but are not essential for it. For example, the α subunit of RNA polymerase becomes adenosine diphosphate—ribosylated a few minutes after infection by the action of either of two phage gene products: gpalt, which is activated during virion assembly and injected along with the DNA, and gpmod. ADP-ribosylation modifies polymerase specificity, but both successful infection and shutoff of host promoters occur in its absence (in *alt/mod* double mutants.) Shutoff of host transcription is also effected by a T4-encoded transcriptional terminator (gpalc) which recognizes local substitution of HMC for C near the terminator site; thus the chemical marking of phage DNA by HMC is used not only in the specific degradation of host DNA but also in reducing transcription before degradation (39). And a T4-encoded inhibitor of σ70 could explain how gp55 supplants σ70 at late times, but again this protein is nonessential for late gene function (55). T4 also has some noteworthy features with no known regulatory role, such as genes that are discontinuous because of intron splicing or ribosomal skipping (6,70).

Assembly

Once late genes are expressed, the stage is set for assembling virions. The T4 virion was the first biological structure of comparable complexity whose pathways of assembly were worked out, from a combination of physical and genetic knowledge. The major technical factors facilitating the study were, first, the availability of conditionally lethal mutations in genes for individual virion proteins; second, the fact that the final steps of the pathway were readily executed *in vitro* so that reactions in the pathway could be distinguished from irrelevant interactions or side reactions that do not contribute to the yield of functional plaque-forming particles; and third, the facility of operations by then standard in phage work, such as assay for complexes stable to dilution.

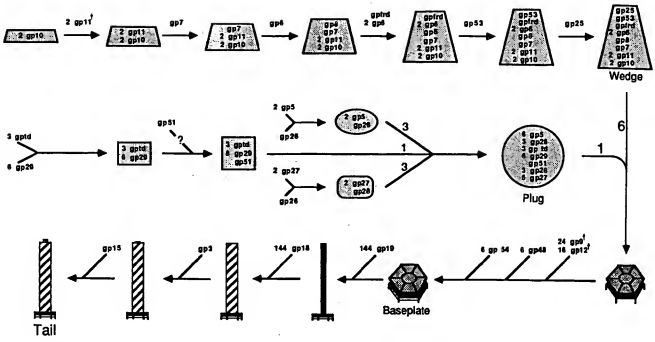


FIG. 1. Assembly pathway of T4 tails. Proteins must be added in the indicated order except that gp11, gp12, and gp9 (indicated by †) can be added at later times as well. From Casjens and Hendrix (14), with permission.

The result of the studies suggested several general principles whose rationalization as effective strategies and extension to other phages has been subsequently assessed and updated (26). One of these is the repeated use of subassembly pathways rather than a single linear pathway. Thus, heads, tails, tail fibers, and soluble protein (gp63) catalyzing tail fiber addition are all made separately. Heads and tails combine first to form complexes that are both visible in the electron microscope and stable to dilution, and finally tail fibers are added to the complex. The tail consists of a baseplate (to which the fibers attach) composed of a central hub surrounded by six wedges. Hub and wedges are made separately, then combined to form the baseplate (Fig. 1). Likewise, in tail fiber assembly, the distal and proximal segments are assembled separately, then combined immediately before addition to the head/tail complex. The most obvious advantage of subassembly pathways is that mistakes in assembly can be rejected for each pathway separately rather than ruining entire virions.

A second general rule is that linear elements are built from the end in (toward their junction with the rest of the virion). In tail assembly, baseplates are made first, then the core and sheath are added to them. Likewise, wedges are formed by adding different protein units successively from one of the outermost units, finishing with the one that interacts with the hub. This strategy ensures that the intermediates of the subassembly pathway cannot compete with the final product for assembly into the virion. The order of assembly of the wedges themselves may be prescribed by allosteric changes on addition to the complex or by the recognition of binding sites that contain elements of more than one of the proteins already present. In fiber addition to head-tail complexes, fibers do not add efficiently to base plates until heads and tails are joined; apparently because the bend in the fibers contacts the head tail junction and interacts with it. A protein (wac) in the collar serves this function *in vitro*, although gpwac is not essential for plaque formation *in vivo*. In free virions, fibers are frequently wrapped around the tail sheath, becoming extended only under specific conditions.

The joining of fibers to head/tail complexes attracted attention early because it proceeds by catalysis of a noncovalent interaction. The catalyst (gp63) also functions as an RNA ligase that apparently acts during (but is not essential for) DNA synthesis. Protein catalysis of a one-step noncovalent interaction was novel at the time, and the discovery was too far in advance of later work on chaperonins to have wide impact.

Tail assembly raised the question of length determination. The core and sheath are both formed by polymerization of single protein species. Core assembly is initiated through nucleation on the baseplate, followed by addition of subunits to build the structure. How is the number of subunits counted, so as to ensure the observed uniformity among virions? The answer (better documented for λ than for T4) is by use of a linear measuring stick protein that is embedded within the core and eliminated when assembly is complete (40).

The T4 tail fiber is a bent linear structure with several specific protein monomers or dimers arranged end to end, terminating in a dimer of gp37. This protein at the tip of the tail fiber is so arranged that its amino acid sequence (and hence its genetic map) is colinear with the fiber axis. This was shown in T2-T4 hybrids recombinant within gene 37, where specific antibodies decorate that part of the structure derived from the phage against which the antibodies are directed (5). The specificity for host cell receptors and neutralizing antibodies lies in the distal portion, near the tip, which is also the location of mutations for altered host range.

The pathway of capsid formation in T4 (as for isometric capsids of any virus) is still understood in less depth than tail synthesis, but several established features are noteworthy. Some of these are better understood for other phages (λ, P22) and inferred for T4. The capsid is synthesized first, and it is filled with DNA later. This fits with the determination of DNA content by head size, but is more directly established by the association of capsids assembled early with DNA made late in infection (47). Second, capsids are not built directly into their final form; rather, capsid proteins associate with an internal scaffolding protein, which fills the interior of the prohead and is later extruded through holes in the protein shell, which expands to its final shape before DNA enters. Both scaffolding protein (gp24) and the major head protein (gp23) have a shape-determining role [inferred from the aberrations in shape caused by certain missense mutations (11)]. Probably both proteins are added to the structure concurrently. How these proteins actually determine the elongated icosahedral shape is unknown, and the intermediate states are no better understood than for any other closed lattice. After the prohead expands and assumes a more obviously icosahedral shape, two decoration proteins, gphoc and gpsoc, are added at regular positions in the lattice; neither protein is essential for formation of functional virions. Third, assembly requires nucleation from one vertex, in which minor capsid proteins, noncapsid proteins, and host chaperonins (groEL) all participate. This unique vertex becomes the portal vertex through which DNA enters the virion and to which the tail later attaches. DNA packaging requires two phage proteins (gp16 and gp17) that probably act at the portal vertex but are not present in the final virion. gp17 has been identified as the endonuclease that cuts the DNA from the concatemer during packaging (26). Another packaging gene, gp49, is needed for processing the DNA substrate. Only linear DNA is packaged, but intracellular T4 DNA becomes highly branched during replication and recombination; gp49 cleaves Holliday junctions and other branched structures, creating packageable lengths of linear DNA (9). Fourth, processing of gp23 takes place concurrently with assembly. Figure 2 presents the general scheme for virion assembly in double-stranded DNA phages.

Lysis

The final stage of the T4 cycle is cellular lysis, releasing virions into the medium. Lysis requires two gene products: a lysozyme (gpe), which attacks the bonds joining *N*-acetylglucosamines in the rigid murein layer, and a holin (gpt), which creates holes in the inner membrane, allowing the lysozyme to reach its substrate. Other genes affect the timing of lysis. The status of gpt as a holin is not fully established and rests largely on analogy to λ gpS (73). In *e t*$^+$ mutants, the cell dies and the membrane potential collapses at the normal time of lysis, whereas in *e*$^+$ *t* mutants, the cell continues to produce lysozyme and intracellular phage long after the normal time of lysis.

It is obviously desirable that the time of lysis be coordinated with the rest of the lytic cycle, but no regulatory system serving this function has been identified. The timing of lysis appears to be determined largely by the accumulation of gpt, formed coordinately with other late proteins because of transcriptional control. One genetic system that regulates time of lysis under certain conditions is the *r*II locus. T4*r* mutants were discovered early by their ability to form larger than normal plaques, and the *r*II mutants have the related property of inability to plate on strains expressing the *rex* genes of λ [which provided the appropriate selective conditions for Benzer's classical studies on fine structure genetics (7)]. However, their effect on lysis and plaque size is attributable to the absence, in *r*II mutants, of lysis inhibition. When a cell infected with T4*r*+ is superinfected before lysis by another T4 particle, lysis is delayed and intracellular phage development continues, up to a period of several hours beyond the normal latent period of 24 min. The rIIA and rIIB membrane proteins must mediate this response and must likewise destroy membrane damage inflicted by the λ *rex* system in response to superinfection, but the mechanism in unknown.

Phage T7

Each group of T phages has its unique properties. A brief account of phage T7 can indicate some of the major differences from the T4 paradigm. The phage T7 virion contains 39,936 bp of linear double-stranded DNA (completely sequenced). It is terminally redundant (like T4) but not cyclically permuted. The length of the terminal overlap is 160 bp (22). Immediately after infection, host RNA polymerase initiates transcription from a promoter near the left end of the molecule, whose products include an antirestriction factor, a protein kinase that phosphorylates and inactivates *E. coli* RNA polymerase, a new T7-specific RNA polymerase, and a DNA ligase. After a few minutes, T7 polymerase replaces host polymerase and all transcription takes place on the 81% of the genome toward the right end. The first genes transcribed by T7 polymerase

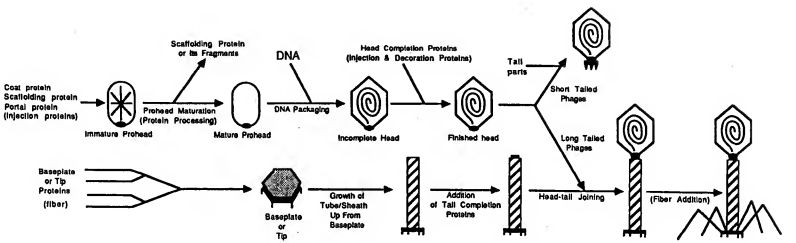


FIG. 2. Generalized assembly pathway for dsDNA phages with tailed virions, such as T4 and λ. From Casjens and Hendrix (14), with permission.

(nearer to the left end) include a DNA polymerase and recombination genes. Later, after replication ensues, the late genes encoding virion components and at least one lysis gene are expressed. Unlike T4, late gene expression does not require replication, and the host polymerase is replaced rather than reprogrammed. Even without inactivation of host polymerase, the T7 RNA polymerase competes so effectively for nucleoside triphosphates that it redirects almost all transcription to T7 promoters, a feature that has been exploited in the design of T7-based expression vectors for DNA cloning.

The first cycle of DNA replication proceeds bidirectionally from a unique origin of the linear monomer and proceeds toward the ends. As with T4, concatemers are probably generated through strand invasion of terminal segments into other molecules within the 160-bp terminal repeat. Replication then generates concatemers in which the unique 39,616 viral sequences alternate with single copies of the 160-bp repeat. If complete virion DNA molecules with repeats at both ends were cut and packaged from such concatemers, half of the genomes would lack terminal repeats. Watson (69) suggested that T4 might solve the problem by cutting the DNA in opposite strands at the two ends of the repeat, creating 160-base protruding 5′ ends, followed by DNA synthesis to produce flush ends. The actual solution is apparently more complicated and wasteful. Replication of one strand initiated from a nick near the 160-bp segment proceeds leftward through the 160-bp segment to a reverse repeat beyond the segment. The displaced strand is then nicked at the end of the reverse repeat, which folds back and initiates synthesis rightward through the 160-bp segment. Thus, two copies of the 160-bp segment are present, which can be trimmed to size and packaged (16). Packaging proceeds from right to left processively along the concatemer (64). Because T7 packages the DNA between two specific sites, the amount of DNA per head is not fixed by head size (as it is in T4). From study of T7 variants with deletions or duplications, the packaging limits of T7 are about 85% to 103% of normal genome length (F.W. Studier, personal communication).

Small DNA Phages

A class of phages package small covalently closed circles of (+) strand genomic DNA in their virions. There are two groups: isometric and filamentous. Although the two groups are not detectably related in sequence and differ in modes of attachment, packaging, and egress, their replication mechanisms are similar.

Isometric Phages

The prototype is ϕ X 174. The virion (25–35 nm) has an icosahedral T = 1 capsid, composed of three polypeptide species: gpF (60 copies), gpG (60 copies), and gpH (12 copies). gpG and gpH are considered spike proteins and can be physically removed without disrupting the integrity of the capsid. The spike proteins are necessary for attachment to cells, and gpH enters the infected cell with the DNA and participates in phage development. Such participation is not essential for all productive infection, because transfection by pure (+) strand DNA from virions leads to some phage production.

Once within the cell, the DNA becomes coated with single-stranded binding protein (Ssb), then begins to replicate. The first cycle [parental single strand → double-stranded replication form (RF)] is mediated entirely by host enzymes. Transcription by host DNA-dependent RNA polymerase at a unique site near a DNA hairpin provides a primer, which is extended by the host replication complex and then digested away. At completion of the cycle, the new DNA strand is ligated to give a covalently closed circle (RF). Transcription from RF by host RNA polymerase leads to production of phage proteins, among them a dimeric protein gpA, which works only in *cis* (i.e., an *A*- mutant cannot be complemented in *trans*, perhaps because the protein binds rapidly and irreversibly to DNA near the site of transcription). Rolling circle replication is initiated when gpA binds to RF at a specific origin site and nicks its parental (+) strand, which is then extended by action of a host-specified DNA helicase (Rep) followed by the host elongation complex. The displaced 5′ end is not free, but remains bound to the growing point through the gpA dimer. The displaced strand is coated with Ssb and remains single stranded (no lagging strand synthesis). When one genome length of DNA has been synthesized and the growing point comes to the origin, the displaced DNA is cut from the newly synthesized DNA and its ends are ligated, giving a single-stranded circle and a double-stranded (RF) circle, both of which can recycle through further rounds of replication. Thus, a fraction of the (+) strands in infected cells are linear and from monomer to dimer length, whereas both + and - strands are present as monomer circles.

With time, virion proteins accumulate and form empty proheads, and concurrently a phage-specified DNA-binding protein (gpC) is made. On some of the + strands, gpC competes with Ssb for binding to the Rep-gpA complex, and these interact with proheads to become encapsidated into virions. Thus, the DNA is directed toward packaging before it encounters the prohead (4). The replicating pool then ceases to expand as + strands are packaged. The cell cycle ends with lysis, mediated by the phage gpE protein. The mode of gpE action is uncertain, but does not involve either a lysozyme or a holin (73).

Filamentous DNA Phages

The virions of filamentous phages, of which phage fd is the prototype, differ from those of most helical viruses

in that the capsid surrounds a DNA molecule that is not linear but circular and therefore is doubled back on itself, although unpaired. The replication cycle of the filamentous DNA phages is similar to that of the isometric phages. Infection is initiated by adsorption of the end of a virion to the terminus of an F-coded pilus, used during bacterial conjugation. The details of DNA entry are obscure. As with ϕ X 174, ssDNA formed at late times is coated with a phage-specified DNA-binding protein (gp5) that directs the DNA to packaging rather than to recycling. Packaging is concomitant with egress. The major coat protein (gp8) is embedded in the cell membrane. As phage DNA enters the membrane, gp5 is exchanged for a helical capsid of gp8. Subunits of gp8 are proteolytically processed as they are added to phage DNA, and virions are extruded through the cell envelope without loss of cell viability. Except for a few enveloped phages, these are the only phages whose egress does not require lysis.

The filamentous phages (especially M13) are popular as cloning vectors because of their high yields, the absence of a packaging limit, and the single-strandedness of the product. Cloning is generally accomplished through cutting and ligation of RF DNA, followed by transfection.

RNA Phages

The common RNA phages have isometric T = 1 virions each containing a single linear single-stranded RNA molecule about 4 kb in length, encoding three to four genes and capable of assuming extensive secondary structure. They enter bacterial cells by attachment to the sides of F pili (rather than the ends, as with the filamentous DNA phages). Like the filamentous phages, they are only able to infect bacteria that harbor the F plasmid. The four genes of phage MS2 encode the coat protein (C), an RNA-dependent RNA polymerase (P), a lysin (L), and a fourth protein (A for *a*ssembly) present in one copy per virion. The genes are arranged in the order 5′ A-C-(L)-P3′, where the L gene overlaps the 3′ end of C and the 5′ end of P.

In these phages, there is no known distinction between replication and transcription. All RNA copies are full length, and gene expression is controlled at the translational level. Nevertheless, they have developed a program of temporal control exquisitely tailored to their needs. The *A* gene at the 5′ end is available for translation only in nascent RNA that has not yet folded into its most stable secondary configuration, a reasonable strategy for obtaining one copy per virion. The *C* gene, on the other hand, is available for translation at all times. The *P* gene is under a more complex control. Translation of the *C* gene is needed to open the secondary structure that otherwise sequesters the *P* ribosome binding site; but coat protein itself, binding to the RNA upstream of the *P* gene, inhibits polymerase translation. The first control delays the onset of phage replication whenever conditions for protein synthesis are so poor that translation initiation of both genes is unlikely within a brief window of time; the second means that, as with T2 or T7, synthesis at late times is diverted to virion components (in this case, coat protein) with minimal competition from polymerase synthesis. The *L* gene lacks a ribosome binding site of its own. Lysin synthesis results from reattachment of ribosomes that initiated at the ribosome binding site for *C*, then dissociated due to high frequency frameshifting and reattached at the AUG for *L* (41). By tying the accumulation of lysin to the rate of coat protein synthesis, this arrangement should help coordinate the time of lysis with the rest of phage development.

In broad outline, the mechanism of RNA synthesis is similar to that of most linear single-stranded RNA viruses. RNA phage replication enjoys the distinction of being the first system in which *in vitro* replication of a nucleic acid was shown to initiate at a specific origin (32). In conjunction with three host proteins (elongation factors Tu and Ts and ribosomal protein S1, all part of the machinery for host protein synthesis), the phage polymerase initiates replication at the 3′ end of a (+) strand. Generally, several subsequent initiations take place at the same end before the first (-) strand is completed. When completed (-) strands dissociate from their (+) strand templates, they are used in a similar manner as templates for (+) strand synthesis.

Because of their small size and rapid mutation rate, RNA phage genomes have been used in studies of *in vitro* evolution, where enzyme is supplied and those RNA molecules structurally capable of the most rapid rate of multiplication are selected (37). The most consistent result was selection for much smaller molecules, which always include the terminal sequences recognized by the polymerase. The results are instructive with respect to the defective interfering particles that accumulate *in vivo* during the high multiplicity passage of many viruses. In absence of added template, the replicase can also create short sequences *de novo* and copy them (10).

TEMPERATE PHAGES

Temperate phages are mostly large (>20 kb) double-stranded DNA viruses. Three prototypes can be distinguished based on their modes of establishment of lysogeny. The first, represented by coliphage λ, insert their DNA into the chromosome at one or a few preferred sites. The second, represented by phage Mu-1, insert anywhere in the host chromosome by use of a phage-coded transposase. In the third, represented by phage P1, prophage DNA is not inserted into the chromosome but instead maintained as a plasmid. Some other phages, such as satellite phage P4, can be maintained either as inserted prophages or as plasmids. The inserted state can be stably maintained only if the phage genes for autonomous replication are repressed, whereas in the plasmid state some viral functions are expressed without killing the cell.

Phage λ

λ development has become a paradigm for gene control in developing systems. An infected cell has one of two options: either to lyse and produce viral progeny or to survive as a lysogen. Once a cell is committed to one of these pathways, there are several safeguards to stabilize the decision.

The λ virion is isometric with a T = 7 configuration of capsomeres and a long tail terminating in a single fiber. The virion DNA is a 48,502-bp linear molecule with complementary 12-base 5′ overhangs. Once injected into the cell, the DNA ends pair and are ligated by host ligase to generate a covalently closed circle.

Commitment

After infection, host RNA polymerase initiates transcription at three promoters: *pL*, *pR*, and *pR′*. The *pL* and *pR* transcripts each contain one gene (*N* and *cro*, respectively). The *pR′* transcript contains no genes. The Cro protein binds to the operators controlling *pL* and *pR* and represses transcription from them; however, because *cro* itself is transcribed from *pR*, the Cro concentration never reaches a level where repression is complete. The gpN protein is an antiterminator, which allows transcription at *pL* and *pR* to read through ordinary transcription signals. The seminal early observations leading to this conclusion were, first, that the λ gene *exo*, which is leftward and downstream of *N*, requires for expression both a functional gpN (suppliable in *trans*) and a derepressed *pL* (in *cis*) and second, that terminators of ordinary bacterial operons such as *trp* are read through when placed downstream of *pL* in the presence of gpN (27,48).

Antitermination by gpN is effected by binding to specific RNA sequences (*nut* sites) that occur early in the transcripts from *pL* and *pR*. The *nutR* RNA includes a binding site (*boxA*) for host protein NusA and a stem-loop sequence (*boxB*) specific for λ's gpN, where NusA is the product of one of the *nus* genes, isolated by screening for mutants with an "*N* undersupply" phenotype, so that even λ*N*$^+$ behaves like an *N*$^-$ mutant on the *nus* mutant strains. Various λ-related phages have distinct *N* specificities determined by the sequence in the *boxB* loop. The gp*N*-NusA complex bound to *nut* RNA interacts with the RNA polymerase that has passed the *nut* site and is transcribing downstream from it and modifies transcription so that normal transcription signals are no longer recognized. The complex is stabilized by interaction with other host factors: NusB, NusG, and ribosomal protein S10 (50).

The result of antitermination is that both leftward and rightward transcription proceed beyond the primary terminators (*t*L1 and *t*R1, respectively.) The leftward transcript includes several genes whose functions are not essential to either lytic or lysogenic development, although some of them play ancillary roles in DNA replication (*exo*, β) or lysogenization (*cIII*). The rightward transcript includes *cII* (whose product promotes lysogenization), replication genes *O* and *P*, and gene *Q*, a positive regulator of late transcription.

The gpcII protein is a positive effector of transcription initiating at a site within *c*II and proceeding leftward beyond *pR* through genes *cI* and *rexAB*. The cI product (repressor) binds to the same operator sites as Cro and represses transcription from *pL* and *pR*, thus stopping all transcription characteristic of the lytic cycle.

If gpcII were expressed at a high rate in all infected cells, they should all become lysogenic. In fact, only some cells do. Although all cells in a culture are under the same ambient conditions, some become committed to lysis and others to lysogeny. The reason for that is still not very well understood, but a great deal is known about the players in the decision and some of the factors influencing the outcome. The gpcII protein works effectively only at high concentration (probably because the active form is a multimer); and it is metabolically unstable so that a high concentration can only be achieved by a high transcription rate (unlike Cro, which accumulates with time). The instability is due to *in vitro* proteolysis; when the host protease most responsible for gpcII degradation (HflA) is eliminated by mutation, the rate of lysogenization by λ (and especially by cIII mutants) increases. The gpcIII protein acts by inhibiting gpcII degradation.

So the commitment to lysogeny is effected by achieving a sufficient gpcII concentration before it is too late. When is it too late? What clock is ticking? The clock is the steady accumulation of Cro, which causes a progressive decrease in the transcription rate from *pR* and therefore in the rate of gpcII synthesis. Once that drops below a critical level, the cell is irrevocably committed to the lytic cycle. The fraction of cells that commits to lysogeny increases with the multiplicity of infection; in fact, under some conditions, the lysogenization frequency of singly infected cells is essentially zero, as might happen if a single copy of the *c*II gene could not work fast enough to build up a high concentration of product.

In those cells where repression is established, further events are needed to ensure stable lysogeny. First, repressor synthesis must continue in the lysogen so that the lytic cycle functions are shut off permanently. Once repression is established, *pL* transcription stops and gpcII disappears. Repressor does not disappear as well, because *cI* transcription now commences from the maintenance promoter *pM* immediately to the left of *oR*.

The *oR* and *oL* operators are both tripartite, with three binding sites separated by short spacers. Both repressor and Cro recognize these binding sites, but with different relative affinities. Cro binds most tightly to *oR*3 (which represses leftward transcription from *pM*) and less tightly to *oR*2 and *oR*1 (repressing rightward transcription from *pR*). Repres-

TABLE 1. *Gene regulation by repressor (R) and Cro (C) at λ oR*

Leftward transcription from *pM* (cI←—)	*o*R1	*o*R2	*o*R3	Rightward transcription from *pR* (cro—→)
Low	—	—	—	On
On	—	R	R	Off
Off	R	R	R	Off
Off	C	—	—	On
Off	C	C	—	Off
Off	C	—	C	Off
Off	C	C	C	Off

Adapted from Ptashne (57) with permission.

sor, on the other hand, binds tightly and cooperatively to *o*R2 and *o*R1, which both represses transcription from *pR* and stimulates transcription from *pM*. At high repressor concentration (unsustainable in the steady state) repressor also binds to *o*R3 and represses its own synthesis (Table 1). Thus, once a high concentration of repressor is achieved through transcription from *pE*, repression is self sustaining (57). The inhibition by Cro of spontaneous initiation at *pM* reinforces the decision to follow the lytic pathway.

Commitment to lysogeny requires not only repression of lytic cycle genes but also insertion of phage DNA into the host chromosome. Insertion is mediated by the phage-coded integrase (Fig. 3). The structural gene for integrase, *int*, is transcribed from two promoters: the major leftward promoter *pL* and a cII-activated promoter *pI*. The latter is the important promoter for lysogenization. Its control by gpcII assures that insertion occurs in those cells where repression is established.

Integrase promotes reciprocal recombination between sites on phage and host DNA, which inserts the phage DNA into the continuity of the host chromosome. The recombination occurs within a 15-bp sequence that is identical in phage and host. Both *in vivo* and *in vitro*, the minimal length of specific sequence required in the *attB* partner is 21 bp and that in the *attP* partner is about 240 bp. A linear

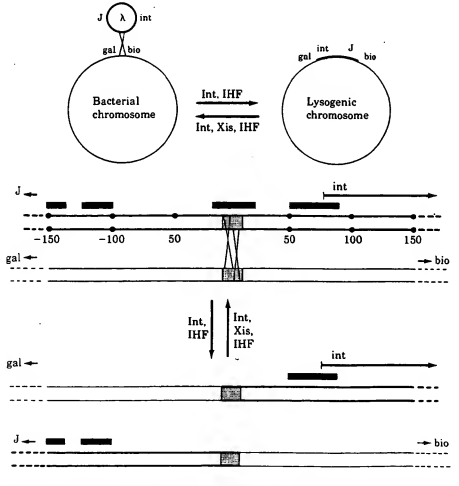


FIG. 3. Insertion of λ DNA into the bacterial chromosome. Top: Overall process. Bottom: Detail of action at the crossover point. The 15 bp identity between phage and host DNA is stippled. The heavy bars overlying the phage are Int-binding sites. The 3′ end of the *int* gene overlap the right site. (Reproduced from 12a with permission.)

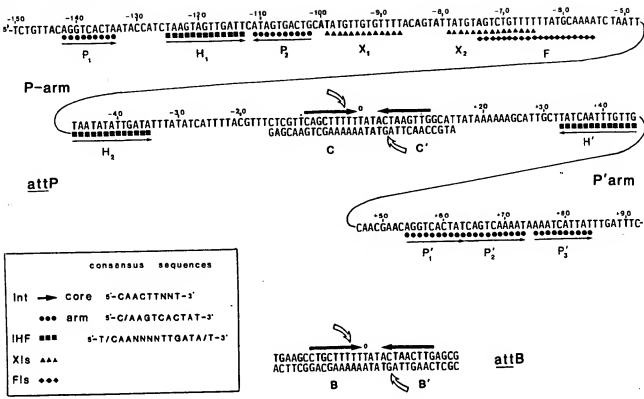


FIG. 4. Sequence of λ *attP* and λ*attB*, showing binding sites for proteins required for insertion and/or excision. P_1,P_2, P_1', P_2',P_3': Arm binding sites for Int. C,C',B,B': Core binding sites for Int. H_1,H_2H': IHF binding sites. X_1,X_2: Xis binding sites Fi: Fis binding site. Curved arrows indicate positions of strand exchange in core. (Reproduced from 42a with permission.)

attB molecule works *in vitro*, but *attP* must be supercoiled for optimal activity.

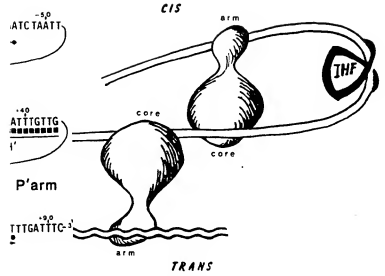
Integrase has two DNA binding sites. The strong binding sites (in the *N*-terminal part of the molecule) recognizes arm sites, which are located on *attP* at position –130, –80, +60, +70, and +80 (where 0 is the center of the 15-bp identity) (Fig. 4). Integrase molecules bound to these sites on DNA, bent appropriately by the host-coded DNA bending protein integration host factor (IHF), are so positioned that they can bind (in their *C*-terminal domains) to core sites, which are symmetrically disposed about the crossover point and approach the consensus 5'CAACTTNNT3' (Fig. 5). Core sites are present in both *attP* and *attB*. The actual crossover then takes place by an exchange on the left between the "top" strands of the two double helices to create a Holliday junction, then branch migration 7 bp to the right, and finally resolution by exchange on the right between the bottom two strands. The left-right orientation here is determined by the positions of the arm sites on the *attP* partner, not by the core sequences themselves (42). Host and phage sites are aligned by protein-protein and DNA-protein interaction, and DNA-DNA recognition is restricted to the 7-bp overlap segment where branch migration occurs.

Commitment to lysogeny is further reinforced by a third cII-activated promoter pAQ, located within gene *Q* reading leftward. This antisense transcription apparently minimizes *Q*-promoted late gene transcription characteristic of the lytic cycle.

The lysis/lysogeny decision can be shifted by various external treatments, such as intracellular cyclic adenosine monophosphate concentration, ion concentration, and multiplicity of infection. Changing these variables can shift the proportion of infected cells that are channeled toward lysogeny rather than lysis. The most noteworthy (and still least understood) feature is how the system is poised so that a significant fraction of the cells almost always goes in each direction, over a wide range of ambient conditions; as though a large degree of indeterminacy were built into the system. One possible contributor to the indeterminacy may be the molecular fluctuations early in infection, such as whether the first molecule of gpN (which promotes lysogenization by increasing *cII* and *cIII* transcription) is made before or after the first molecule of Cro.

Lysogeny

Once established, the lysogenic state is quite stable. Lysogenic cells divide to give lysogenic progeny. Lysogeny can break down either through spontaneous switching to the lytic cycle (which happens about once every 10^4 cell divisions in



3 5. DNA looping by Int. The arm and core sites binding an monomer may either lie on the same partner (cis) or opsite partners (trans). DNA bending by IHF brings the cis as into favorable juxtaposition. The exact configuration of integration complex is unknown. (Reproduced from 42a h permission.)

actively growing culture) or by spontaneous loss of phage with cell survival (which happens less than once ry 10^6 cell divisions). As long as repressor is present, *cI* scription from *pM* continues, and more repressor is made. epressor concentration rises so high as to saturate *o*R3, n repressor synthesis ceases until it returns to a lower apparently el. Likewise, once the prophage is stably inserted into the on characteromosome, it remains so; there is a low rate of integrase thesis from the *pI* promoter (not activated by gpcII) but ifted by vaentially no excisionase (a phage-coded protein required cyclic adenexcision, though not for insertion).

ıtration, and ome ancillary factors reinforce this stability. One of ariables can se is the need for antitermination by gpN to initiate lytic channeled toelopment. This requirement puts a double lock on the eworthy (andem, because any transcripts made during a momentary stem is poisslocking of the operator sites will not get anywhere until most always gpN that has been made by the first transcript can exnbient condid other transcripts. Because gpN is unstable and acci- cy were built's should be rare, this will seldom happen. Stabiliza- the indeterm of lysogeny is one reason that has been suggested for in infection,use of antitermination by λ. Poising the lysis/lysogeny ich promotesision to depend on molecular fluctuations may be anscription) is ar. Explanations for antitermination have been sought). ause it seems simpler a priori to eliminate the termi- rs rather than to maintain and override them.

ape from Lysogeny

quite stable.

eny. Lysogerbout once every 10^4 cell divisions, stability breaks down. witching to thiled knowledge of what happens in those rare cells is un- 10^4 cell divisilable, but the rate is strongly depressed in *recA*⁻ hosts.

This fact suggests that the rare spontaneous transitions to lytic development have the same primary basis as the mass induction of this transition by exposure to ultraviolet light or other DNA-damaging agents. In that case, the primary event is the activation of the proteolytic activity of RecA protein by products of DNA degradation, such as single-stranded DNA. RecA is not a conventional protease, but it accelerates the specific cleavage of the host LexA protein, which can autocatalyze the same cleavage under extreme conditions. LexA is a repressor for many DNA repair functions, and RecA-promoted LexA cleavage is the initial step in the SOS response that turns these genes on. Like LexA, λ repressor is sensitive to RecA-enhanced autoproteolysis (46).

Turning on the SOS system is one means of inducing the switch to lysis in almost all cells of a lysogenic culture. Another method makes use of λ *cI*ts mutants, which form a thermolabile repressor. With an appropriate mutant, lysogeny is quite stable at 30°C, but a shift to 43°C rapidly inactivates repressor in all cells. After derepression, transcription initiates from *pL* and *pR*. One early consequence of *pL* transcription is excision of the prophage from the chromosome. Whereas insertion requires only one phage-coded protein, integrase, excision requires a second protein, excisionase, as well. Excisionase is encoded by a gene *xis*, upstream from *int* and overlapping it by 20 bp. The cII-dependent *pI* promoter used during lysogenization lies at the 5′ end of *xis*, and the start site is within *xis* so that the *pI* transcript makes only integrase. From the *pL* transcript, *int* and *xis* are expressed coordinately.

The mechanism of excisionase (Xis) action is not fully understood. Excisionase binds to two specific sites in the left arm of *attR*, bends the DNA, and interacts with integrase; one of the two excisionase molecules can be replaced by a host DNA-bending protein, Fis. Conventional enzymes are pure catalysts, and the nature of the catalysts cannot determine the direction of the reaction. Superficially, excisionase appears to reverse the integrase reaction; however, in excision as in insertion, branch migration proceeds from left to right; therefore, excision is not a true reversal of insertion. What remains unsolved is why integrase alone fails to promote the true reversal.

The *pL* transcript is subject to posttranscriptional control. In infected cells, either before commitment or after commitment to lysis, *int* and *xis* are cotranscribed; however, little integrase is made. This is because of exonucleolytic (3′ → 5′) degradation of the transcript after cleavage by host enzyme RNAse III at a site (*sib*) downstream of *att*. This process (retroregulation) prevents untimely or wasteful synthesis of integrase. Retroregulation does not affect *pL* transcripts made from a derepressed prophage because the *sib* site is at the other end of the prophage. It also does not affect *pI* transcripts, which terminate within *sib* and therefore are resistant to RNAse III. Because the *pL* transcript is antiterminated, it reads through *sib* and is degraded; this is another function for antitermination in

the λ life cycle. Despite the elegance and apparent utility of the retroregulation mechanism, deletion or mutation of the *sib* site has no observed effect on λ development as normally studied in the laboratory.

The potentiality of controlling insertion and excision differentially provides a rationale for the use of arm sites by λ integrase. Some related site-specific recombinases recognize only core sites, but in those cases there is no recognition of directionality.

When lytic development is induced in a lysogen, a small fraction (about 10^{-5}) of the particles produced result from abnormal excision that follows from breakage and joining of DNA at sites other than the *att* sites. If the excised DNA includes the *cos* sites and is within the packaging limits of the λ virion, it can be packaged into virions and used to infect other cells. Thus, a lysate produced by induction can transfer host genes that lie close to the λ insertion site (such as *gal* or *bio*) into recipient cells. This specialized transduction is a form of natural cloning. The transduced bacterial DNA segment need not replace its homolog in the recipient; instead the cell becomes lysogenic for the specialized transducing phage and is diploid for the transduced segment (Figs. 6 and 7).

The genesis of specialized transducing phages may be relevant to the acquisition of oncogenes by retroviruses. Although the mechanisms may differ substantially, specialized transduction has provided a useful model because the entire process can be followed in the laboratory. With retroviruses, until very recently the critical events have generally occurred in nature from inferred progenitors.

Lytic Development

Lytic development can be studied either in infected cells or in derepressed lysogens. Because of the high degree of synchrony obtainable, the method of choice has usually been thermal induction of lysogens carrying *cI ts* prophages. The genetic program and its readout are well understood. At early times, most transcription is from *pL* and *pR*. The only protein of the *pL* transcript that is essential for lytic development is gpN. The *p*R transcript contains three important genes: *O* and *P* (replication) and *Q* (late gene activation). As time proceeds, gpQ accumulates and causes antitermination of the short transcript starting at *p*R′ (downstream from *Q*). The mechanism of antitermination differs from that performed by gpN in that the recognition elements for antitermination are immediately downstream from the promoter and do not cause termination when further displaced from it (29). *Q* antitermination also requires a pause site close to the start site, whereas the role of pausing in *N*-antitermination is less certain.

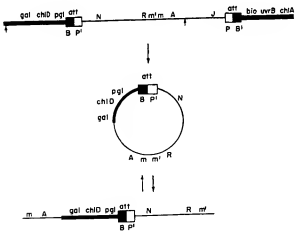


FIG. 6. Genesis of λ *gal* specialized transducing phage. Arrows indicate positions of breaking and joining of heterologous DNA in a particular isolate. m and m′ indicate the ends of mature λ DNA (equivalent to *cos*). (Reproduced from 12a with permission.)

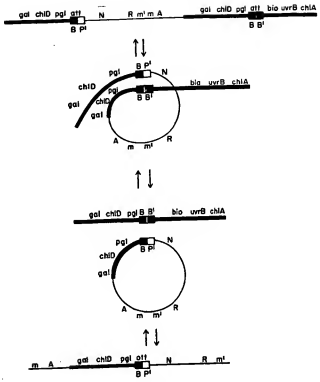


FIG. 7. Lysogenization by λ gal (bottom line) to produce a heterogenote with two copies of *gal* separated by λ DNA. (Reproduced from 12a with permission.)

At any rate, as gpQ accumulates, transcription of genes for lysis and virion formation becomes increasingly rapid. Concurrently, the rate of transcription from *p*L and *p*R decreases as Cro accumulates, producing an orderly transition from early gene expression to late gene expression.

Bidirectional DNA replication initiates from an origin within gene *O*. From studies of hybrid phages with different initiation specificities, it is clear that gpO recognizes the origin and also recognizes gpP, which in turn interacts with the host DnaB helicase. Disassembly of the initiation complex to leave an elongation complex requires

three host heat shock proteins (DnaJ, DnaK, GrpE) not needed for host DNA synthesis. Elongation proceeds, as in the host, by the combined action of DnaB helicase, DnaG primase, and Pol III holoenzyme, the polymerase. Theta-form replication can reproduce monomer circles, but a switch to rolling circle replication occurs in some molecules. λ rolling circle replication (which predominates at late times) differs from that of ϕX in two major respects. First, the tails are double stranded rather than single stranded because lagging strand synthesis takes place on the displaced strand. Second, tails can grow to multigenomic lengths rather than stopping when a single genome has been spun off.

Multigenomic tails are the primary substrates for λ packaging. Empty proheads are first assembled, composed mainly of the major capsid protein gpE. Later, DNA enters the prohead, increases in size and becomes more conspicuously icosahedral, and the smaller decoration protein gpD is added to the shell, with gpD hexamers interspersed among gpE trimers in a T = 7 arrangement.

Efficient DNA packaging requires the presence of specific sequences (called *cos* for *co*hesive *s*ite") at both ends of the packaged genome; hence, multigenomic tails or multimeric circles are good packaging substrates and monomer circles are poor ones. Cutting at *cos* is effected by a heterodimeric protein, terminase, whose subunits are encoded by λ genes Nul and *A*. DNA molecules are packaged from left to right so a DNA monomer can be cut from the concatemer at its left end early in packaging and at its right end after the DNA has entered the virion. The *cos* site (about 200 bp in length) has three components (from right to left): a binding site (*cos B*), recognized by gpNul before the DNA is cut at either end; a nicking site (*cos N*), at which DNA is cut at sites 12 bp apart in the two strands to generate the packaged single strand overhangs of virion DNA; and a termination site (*cos Q*), required for the final nicking at the right end (18). Packaging is processive. After one genome is packaged, the next genome to the right can be packaged with no requirement for recognition of *cos B* on the second or subsequent genomes. This was demonstrated in the packaging of adjacent prophages with different *cosB* specificities determined by λ-related phages (24).

Because λ packages the DNA between two *cos* sites, the amount of DNA per head can be changed by deletions or insertions (as in T7). The packaging limits are from about 79% to 110% of normal λ length. Internal *cos* sites artificially placed within λ at less than or near the lower limit can be packaged uncut, as though scanning for *cos* sites becomes more efficient as the head fills.

Table 2 indicates the packaging limits for some phages described in this chapter. Those with headful packaging are listed with the same upper and lower limits. In fact, there is some random variation around the mean value. Alternative forms with different head sizes, such as the isometric and giant particles of T4 or the λ virions formed by *D*⁻ mutants are not considered.

Genome Organization

The functional clustering of genes on the λ map is notable. Not only do genes acting in a single pathway (head formation, tail formation, replication) lie together on the map, but genes for DNA binding lie close to their target sites, products of adjacent genes in a cluster frequently interact directly, and even within genes determinants are so ordered as to lie close to the elements with which they interact. For example, the *cosB* site is at the right end of *cos*, next to the Nul gene whose product recognizes *cosB*; Nul in turn lies next to *A*, whose product forms the terminase heterodimer with gpNul. In both head and tail clusters, the same tendency is observed, the tail fiber gene *J* being most distal from the head cluster. The *att* site is adjacent to the *int* gene, and *cI* and *cro* are close to their target sites. The N-terminus of gpO recognizes the *ori* site that lies within the 5′ end of *O*, whereas the C-terminus interacts with the product of the adjacent *P* gene. The clustering of major functional groups could increase regulatory efficiency, but the finer orderings may be more related to the disruptive effect of recombination on co-evolved functions.

Phage Mu-1 as a Model Transposon

Mu-1 first attracted attention as a phage capable of inducing bacterial mutations at a high rate. The mutations

TABLE 2. *DNA cutting and packaging*

Phage	Packaging substrate	First cut	Second cut	Packaging limits (% of genome)
T4	Concatemer	Random	Headful	102
P22	Concatemer	*pac* site	Headful	109
λ	Concatemer	*cos* site	*cos* site	79–110
T7	Concatemer	Terminal overlap	Terminal repeat	85–103
ϕX174	Dimer ss of rolling circle	Replication origin	Replication origin	Limited
M13	Dimer ss of rolling circle	Replication origin	Replication origin	Unlimited
Mu-1	Inserted monomer	Near left end	Headful	105
P1	Concatemer	*pac* site	Headful	110

proved to be insertions of Mu-1 into random sites on the host chromosome, disrupting genes or operons into which it inserted.

Like λ, Mu-1 can follow two alternative developmental pathways, leading on the one hand to a transcriptionally quiescent prophage and on the other to a lytic cycle with replication, packaging, and lysis. In λ, the DNA transactions in these two pathways are distinct. Insertion plays no role in λ replication, which proceeds normally if either the *int* gene or the *att* sites are deleted. With Mu-1, on the other hand, replication and insertion are minor variations of a common pathway, transposition. In Mu-1 replication, transposition happens many times during each cycle of infection, which has made it a system of choice for studying the biochemistry of transposition (compared with most bacterial transposons, which transpose naturally at rates such as 10^{-6} per cell division).

Mu-1 DNA is inserted into host DNA at all stages of its life cycle, even in the virion. When linear DNA from the virion is injected into a cell and the Mu-specific transposase is made, the first step in replication is transposition of Mu-1 DNA into the chromosome. Among bacterial transposons, the transposition mechanism may either be replicative (generating transposon copies in both the donor and the target site) or conservative (excising the transposon from the donor and inserting it into the target site). At least in cells destined to become lysogenic, the initial transposition is conservative rather than replicative. In cells destined to lyse, Mu-1 undergoes repeated rounds of replicative transposition.

The initial steps of replicative and conservative transposition are the same. Target sites are cut to generate protruding 5-base 5′ ends. The two donor strands are then cut at the 3′ ends of the transposon and ligated to the protruding 5′ ends of the target DNA. Replication then initiates from the recessed 3′ ends of the target DNA. In conservative transposition, the nicked donor DNA adjacent to the transposon is cut in the other strand and digested away; replication fills the gap, and ligation incorporates the parental transposon bodily into the target site (Fig. 8). In replicative transposition, the donor strands remain attached to the 5′ transposon ends, and replication proceeds through the transposon with ligation to the free 5′ donor ends at the opposite end of the transposon. This produces two semiconserved transposon copies, each joined to donor DNA at one end and target DNA at the other (51). Replicative transposition from a chromosomal donor produces two inserted copies accompanied by chromosomal rearrangement (inversion or excision); if replicative transposition happened from a linear infecting molecule (unaccompanied by subsequent recombination), it would linearize the chromosome.

In Mu-1 packaging, the left end of the inserted DNA is recognized, and cutting takes place within the flanking DNA, 50 to 150 bp to the left of the left end. Packaging is then by headfuls, and the headful length is longer than the genome length. Therefore 500 to 3,000 bp of DNA on the right end of the virion molecule is host DNA. So virion DNA molecules have random segments of host DNA at their termini. When virion DNA is separated into single strands and reannealed, the host sequences are visibly unpaired when viewed in the electron microscope.

An internal segment of Mu-1 DNA about 3 kb in length is invertible through the action of a phage-coded integrase, so reannealed molecules are sometime unpaired in this segment as well. The gene for the tail fiber protein spans one boundary of the invertible segments so that two types of tail fibers with different host specificities can be formed, depending on the orientation. Thus, if Mu grown on *E. coli* K-12 encounters an alternative host such as *Citrobacter*, the rare particle that attaches to *Citrobacter* can invade and replicate. The rate of inversion is about 10^{-2} per generation. Certain other phages, such as P1, have related invertible segments.

In lysogens, inserted Mu-1 can occasionally transpose without induction of a full lytic cycle. Thus, inserted Mu-1 behaves in all respects as a bacterial transposon.

Phage P1 as a Model Plasmid

The P1 virion contains about 100 kb of linear double-stranded DNA in an isometric T = 16 capsid. Like T4, the DNA is cyclically permuted and terminally redundant. Unlike T4 (but like *Salmonella* phage P22, a relative of λ), processive packaging does not initiate at random but instead at specific *pac* sites. After infection, reciprocal recombination within the terminal overlap creates a circular DNA molecule of about 90 kb. Some of this recombination is mediated by a site-specific recombinase (Cre) whose target site (*lox*) is close to the *pac* site.

The cell can now enter either a lytic cycle or a lysogenic cycle. Unlike λ or Mu-1, the lysogenic alternative entails replication of the circular P1 DNA as a plasmid and therefore requires expression of phage-coded replication genes. In lysogens, P1 is maintained at low copy number (about one per chromosome) and has an efficient partitioning mechanism so that both daughters of a dividing cell receive a P1 copy. The mechanism is similar to that used by other stringently controlled low copy number plasmids, such as the fertility factor F, and in some respects to that of the bacterial chromosome. The partitioning mechanism could potentially become ineffective if homologous recombination created a dimeric molecule, which could not pass into both daughter cells; this problem is greatly diminished by the site-specific *lox*-Cre recombination, which rapidly and reversibly equilibrates monomers with dimers. The elements for replication and partitioning (except *lox*-Cre) are localized in 1.5 kb of P1 DNA that can replicate as a plasmid when circularized; it includes a specific replication origin (*ori*R), its cognate replication initiator (*repA*) and a DNA sequence (*incA*) that controls replication by binding and sequestering *repA* (72).

In cells destined to lyse, P1 DNA replication (like λs) proceeds both in a theta form and rolling circle mode, the latter providing the packaging substrate. The replication origin used is distinct from the plasmid origin *oriR*. Rare mistakes lead to occasional packaging of random fragments of host DNA instead of phage DNA. Such fragments are responsible for the generalized transduction of bacterial genes. Generalized transduction at detectable frequencies is observed for most phages that use headful packaging but not by wild-type λ, probably because in λ both cutting (at both ends) and injection require the specific *cosN* sequence, whereas in P1 or P22 initial cutting at sequences resembling *pac* sites can initiate processive packaging. Mutations that relax the specificity and thereby increase the frequency of transduction have been isolated in both P22 and P1 (61,68).

DEFECTIVE PHAGES AND PHAGELIKE OBJECTS

The fact that Mu-1 can behave as a transposon and P1 as a plasmid underscores some of the commonality between phages and nonviral elements.

Transposable elements are common in bacteria, and most of them seem to use a biochemistry similar to that of Mu-1. Most bacterial transposons are smaller than Mu-1, going down to simple insertion sequences a few hundred base pairs long; and typical transposons carry genes affecting conspicuous bacterial traits such as antibiotic resistance. Conservative transposition is the usual mode for some elements, replicative transposition for others. Those using replicative transposition frequently resolve the resulting cointegrate structures through use of a site-specific recombinase that acts on a sequence internal to the element. Some transposons integrate and excise by a recombination mechanism similar in some respects to λ insertion (62).

Bacterial plasmids fall into two general groups: elements with small genomes, present in many copies per cell and replicating randomly with respect to the cell division cycle; and larger elements similar in size to phage P1, present in one or a few copies per cell and partitioned regularly at cell division. Conjugative plasmids resemble viruses as autonomous elements capable of infecting cells, but differ from them in the absence of an extracellular phase.

Paper scenarios in which phages have evolved into or from transposons or plasmids are easy to construct, and laboratory evolution from temperate phage to obligate plasmid or chromosomal element is readily achieved. The present sequence database does not provide a meaningful phylogeny placing various phages among their nonviral counterparts any more than the various phage groups can be clearly related to one another.

DNA sequences closely related to those of temperate phages are found in the genomes of many natural bacteria. For example, Southern hybridization of enteric bacteria with λ DNA usually turns up some λ-related sequences (58). Most of these are probably defective prophages: remnants of previous lysogenization that have since lost phage function through deletion or mutation. The K-12 strain of *E. coli* is naturally lysogenic for λ but also harbors four λ-related elements: DLP12, Rac, e14, and Qin. Such elements could in principle be of some value to the host or could even be precursors rather than descendants of complete phages; however, each defective prophage contains an array of genes from different functional clusters of the phage genome, frequently in the same order found in active phages, so that a role as host elements seems contrived. Defective lysogens arise from active lysogens in the laboratory by prophage mutations and frequently retain superinfection immunity, SOS-induced lysis, and production of structures resembling phage heads, tails, or complete virions.

Many natural strains liberate (or produce on induction) such phagelike or phage-related particles, which have not been shown to be infectious. Many bacteria also produce proteins (bacteriocins) that are toxic to other strains of the same bacterial species but not to the producing strain. Bacteriocins are sometimes encoded on plasmids and act through diverse mechanisms. Operationally, some phagelike objects qualify as bacteriocins.

For example, many natural strains of *Bacillus subtilis* and several other *Bacillus* species harbor a prophagelike

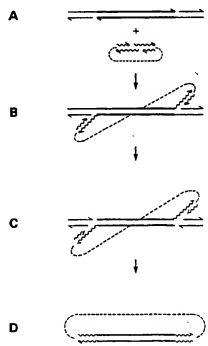


FIG 8. Conservative transposition by Mu-1. **A:** Linear donor molecule nicked at the 3′ ends of Mu-1 strands, plus circular target molecule cut with recessed 3′ ends. **B:** Protruding 5′ ends of target DNA are ligated to 3′ ends of Mu-1. **C:** Donor strands are nicked at 5′ Mu-1 ends (and probably degraded). **D:** DNA synthesis from 3′ target ends fills gaps, ending with ligation to 5′ Mu-1 ends. (Reproduced from 51 with permission.)

element Phage of *Bacillus subtilis,* (PBS), that has been localized on the chromosome. After SOS induction, the cells lyse and liberate many virionlike particles with tails and DNA-filled heads. This is not simply a phage whose sensitive host remains to be discovered because the large majority of particles contain random segments of host DNA with no significant preference for packaging prophage DNA. Like a bacteriocin, these particles kill other *Bacillus* cells. The producing strain is resistant, at least sometimes because the particles fail to be adsorbed. There are several varieties of PBS, each with a different killing specificity (74).

PBS has some similarity to mutants of generalized transducing phages that have lost (in this case, completely) their extreme preference for initiating packaging at phage-specified *pac* sites. Its wide distribution and the variety of killing specificities suggest that the bacteriocinlike activity confers a selective advantage to the host. Inasmuch as some bacteriocins are single protein molecules, a DNA-filled virion seems an especially clumsy and expensive kind of bacteriocin to have been so successful in evolution.

EVOLUTION AND NATURAL BIOLOGY OF PHAGES

For historical reasons, the coliphages have become workhorses of molecular biology and remain among the most thoroughly understood biological objects. Phages are found throughout the prokaryotic world in a variety comparable with that of eukaryotic viruses, including enveloped virions with double-stranded RNA genomes, (67), linear DNA genomes with protein covalently bound to their 5′ ends (59), and satellite viruses like P4 [which forms virions with the capsid protein of its helper virus P2, induces late gene function from P2 prophage, lysogenizes P2 lysogens using its own integrase but can establish a plasmid state in P2 nonlysogens, and represses its own genes by the use of antisense RNA rather than a protein repressor (44)]. Compared with our knowledge of phage development in infected cells, little is known about their natural ecology or population biology. Some aspects are worthy of note.

Abundance

In some natural environments phagelike particles are much more numerous than one might suspect from plaque assays on susceptible hosts. In estuaries, concentrations as high as 10^8 have been reported. If most of these are infectious (and they may very well not be), they potentially could play a major role in the natural cycling of the biosphere (30).

Host Defense Mechanisms

Bacteria have evolved mechanisms for phage resistance, and phages have evolved counterstrategies, presumably indicating in some measure the significant impact that phage infection can have. A host defense mechanism unique to prokaryotes is restriction/modification. Restriction enzymes damage DNA that has not undergone chemical modification at specific sites (most frequently methylation) by cleaving it either at those sites or nearby; nonspecific exonucleases then degrade the resulting linear fragments. Cells protect their own DNA from the restriction enzymes they make by modifying it. A phage that escapes restriction becomes modified and is fully infectious on a restricting host in the next cycle; so the strategy can only be effective for the host if the strains sensitive to a given phage have various restriction specificities. A strain with a rare restriction specificity will exclude almost all phage it encounters and will be favored over one with a common specificity; so eventually the rare specificity will become common. The recognition sites for restriction enzymes are generally symmetrical, even if the cleavage takes place at a variable distance from the recognition site; this allows the modification of either strand to prevent cleavage, which assures that, during semiconservative replication, a new strand that has not yet been modified is protected from destruction.

Phages have many strategies for avoiding restriction. The first level is a paucity of recognition sequence. Phages of *Bacillus*, where many known restriction enzymes recognize 4-base palindromes, have few such palindromes in their DNA; coliphages are deficient in 6-base palindromes, which are sites for some of the enzymes they naturally encounter (38). Various phages also have enzymes that destroy 5-adenosylmethionine, a cofactor for restriction (T3), expedite modification of phage DNA before restriction enzymes can cleave (λ), or block recognition sites by wholesale nonspecific DNA modification (T4).

Some restriction enzymes are encoded by chromosomal genes, but many are encoded by phages or plasmids. Many resistance mechanisms contributing to defense of the host may in fact be manifestations of competition between extrachromosomal elements preserving their cellular territories.

One element may exclude others by diverse mechanisms, and in many cases exclusion may be an incidental side effect of the phage life cycle. The repressors of temperate prophages (essential for the internal stability of lysogens) render the lysogens immune to superinfection by phage of the carried type, and this superinfection immunity probably selects for diversification of repressor specificities. Certain phages, such as the ϵ phages of *Salmonella*, encode enzymes that alter the polysaccharide surface receptors used in phage attachment, so that lysogens are not only immune but also cannot be penetrated by DNA of the same type. The alteration also takes place during the lytic cycle and reduces phage loss from attachment to fragments of cell envelope after lysis. The λ-related phage HKO22 has a variant of the λ*N* gene that competes with λN when λ infects an HKO22 lysogen and causes termination rather than

antitermination of early λ transcripts (54). The λ Rex proteins trigger a collapse of membrane potential when a λ lysogen is infected with certain other phages (such as T4 rII mutants), preventing the production of viral progeny (56). These last two mechanisms do not protect the infected cell but may protect other cells in a colony by impeding viral replication.

Natural Recombination

Third, the study of natural isolates has provided some perspectives on the degree of variation and the frequency of recombination among phages in nature. The λ-related (lambdoid) phages are a good example.

A note on the way that phage workers have approached taxonomy may be useful because it differs substantially from the method used by many medical virologists. In medical virology, once a virus such as measles is identified as a disease agent, many isolates from different clinical sources are examined, and those sharing some acceptable degree of similarity to the type of virus are classified as measles virus. For epidemiology, some such approach is necessary. Phage workers are not epidemiologists, and their interest in natural variation is relatively recent. At least from the Luria-Delbrück era onward, investigators have been more interested in developing highly defined experimental systems. To that end, each isolate is given its own name. "λ" refers only to the prophage present in *E. coli* K-12 and its laboratory descendants; "K-12" refers to one bacterial isolate, not to other natural strains with similar properties. Workers with a different perspective might have classified all lambdoid phages as natural variants of λ, and the same can be said of other phage groups such as T2, T4, and T6.

The lambdoid phages have a common genetic map and the ability to generate viable hybrids when crossed in the laboratory. They have been isolated from various sources (usually from lysogenic bacteria) isolated over the past 70 years from Europe, Asia, and the United States). Genetic variation among the lambdoid phages is apparent both from functional specificity and DNA sequence. For example, phages with at least a dozen repressor specificities have been isolated; where the sequences are known, homologies between heterospecific repressor genes are barely discernible. Phages with different repressors have corresponding differences in their *cro* genes and operator sites.

If the λ sequence is compared with that of any other lambdoid phage (or, equivalently, if heteroduplexes of the two phages are constructed) the general result is that the sequences match closely in some parts of the genome, whereas homologies are weak or absent in others. This is not because some sequences are highly conserved, because another phage pair shows a different set of matching segments; instead, the whole pattern strongly indicates that it is generated by frequent natural recombination so that in any phage pair, some portions of their genomes are of recent common ancestry. When defective prophages are compared with known lambdoid phages, they likewise are closely related to different phages in different segments. For example, the defective prophage DLP12 has an *int* gene and a partially deleted *xis* gene related to phage P22, followed by segments homologous to λs *exo* and *P/ren* genes, followed by analogs (*qsr'*) of λ *Q S R* unrelated to any known λ phage, followed by *cos* DNA similar to λs (45). Thus, all the lambdoid phages, including the defective prophages, seem to be drawn from a common gene pool.

Among the lambdoid phages, some genes that may serve viral functions have been borrowed from outside the λ pool. Gene 12 of P22 occupies the same position in the genome as λ gene *P*, but gene 12 is a homolog of the bacterial *dnaB* gene, and λ gpP recruits the host DnaB helicase to participate in phage replication. Wildtype λ has a gene for side tail fibers (*stf*) that facilitate phage attachment. (In the λ commonly used in the laboratories, *stf* has been inactivated by a frameshift.) The *stf* gene is closely related to tail fiber genes of other groups of coliphages (31). Most lambdoid phages encode a true lysozyme that hydrolyzes glycosydic bonds; λ makes instead a transglycosylase that attacks the same bonds but has no detectable homology to lysozyme (73). The λ lysozyme may have been appropriated from some external source, as yet unidentified.

The natural function of genetic recombination is not obvious. In λ as in sexual eukaryotes, it is clear that recombination has a substantial evolutionary impact. But that does not tell us at what level natural selection acts to preserve recombination as a process or whether recombination is an accidental by-product of gene activities selected to function in repair or replication. One classical argument for the value of recombination is that an asexual line deteriorates through accumulation of deleterious mutations, which can be corrected by recombination with other lines that are wild type for the loci of those particular mutations (52). The first experimental demonstration of this Muller's ratchet effect came from phage φ6 (15); corroborative evidence from animal viruses has since appeared (20). Muller's ratchet may be central to the evolution of certain human viruses such as influenza, where the clonally selected epidemic strains are eventually replaced by reassortants with avian viruses (25).

In their natural ecology, phages exhibit many of the same features observed in eukaryotic viruses and will probably find increasing use as model systems.

REFERENCES

1. Anderson TF. The reactions of bacterial viruses with their host cells. *Botan Rev* 1949;15:464–505.
2. Anderson TF, Doermann AH. The intracellular growth of bacteriophages II. The growth of T3 studied by sonic disintegration and by T6-cyanide lysis of infected cells. *J Gen Physiol* 1952;35:657–667.
3. Andrake M, Guild N, Hsu T, Gold L, Tuerk C, Karam J. DNA polymerase of bacteriophage T4 is an autogenous translational repressor. *Proc Natl Acad Sci USA* 1988;85:7942–7946.

4. Aoyama A, Hamatake RK, Hayashi M. *in vitro* synthesis of bacteriophage φX174 by purified components. *Proc Natl Acad Sci USA* 1983;80:4195–4199.
5. Beckendorf SK. Structure of the distal half of the bacteriophage T4 tail fiber, *J Mol Biol* 1973;73:37–53.
6. Belfort M, Ehrenman K, Chandry PS. Genetic and molecular analysis of RNA splicing in *Escherichia coli*. *Methods Enzymol* 1990; 181:521–539.
7. Benzer S. The elementary units of heredity, In: McElroy WD, Glass B, eds. *The chemical basis of heredity*. Baltimore: Johns Hopkins Press; 1957:76–93.
8. Benzer S. On the topology of the genetic fine structure. *Proc Natl Acad Sci USA* 1959;45:1607–1620.
9. Bhattacharyya SP, Rao VB. A novel terminase activity associated with the DNA packaging protein gp17 of bacteriophage T4. *Virology* 1993;196:34–44.
10. Biebricher CK, Eigen M, McCaskill JS. Template-directed and template-free RNA synthesis by Qβ replicase. *J Mol Biol* 1993;231:174–179.
11. Black L, Showe M. Morphogenesis of the T4 head. In: Mathews C, Kutter E, Mosig G, Berget P, eds. *Bacteriophage T4* Washington, DC: ASM Publications; 1983:219–245.
12. Campbell A. Sensitive mutants of bacteriophage I. *Virology* 1961; 14:22–32.

12a. Campbell A. Genetic structure. In: Hershey AD, ed. *The bacteriophage lambda*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; 1971;13–44.

12b. Campbell A. Transposons and their evolutionary significance . In: Nei M, Koehn RK, eds. *Evolution of genes and proteins*. Sunderland, MA: Sinauer.

13. Campbell AM. Episomes. *Adv Genet* 1962;11:101–146.
14. Casjens S, Hendrix R. Control mechanisms in dsDNA bacteriophage assembly. In: Calendar R, ed. *The bacteriophages*. New York: Plenum; 1988:15–91.
15. Chao L. Fitness of RNA virus decreased by Muller's ratchet. *Nature* 1990;348:454–455.
16. Chung Y-B, Nardone C, Hinkle DC. Bacteriophage T7 packaging. III. A "hairpin" end formed on T7 concatemers may be an intermediate in the primary reaction. *J Mol Biol* 1990;216:939–948.
17. Cohen SS. Virus-induced acquisition of metabolic function. *Fed Proc* 1961;20:641–649.
18. Cue D, Feiss M. A site required for termination of packaging of the phage λ chromosome. *Proc Natl Acad Sci USA* 1993;90:9290–9294.
19. Doermann AH, Eisenberg FA, Boehner L. Genetic control of capsid length in bacteriophage T4. I. Isolation and preliminary description of four new mutants. *J Virol* 1973;12;374–385.
20. Duarte EA, Clarke DK, Moya A, Elena SF, Domingo E, Holland J. Many-trillionfold amplification of single RNA virus particles fails to overcome the Muller's ratchet effect. *J Virol* 1993;67:3620–3623.
21. Dulbecco R. Production of plaques in monolayer tissue cultures by single particles of an animal virus. *Proc Natl Acad Sci USA* 1942; 38:747–752.
22. Dunn J, Studier W. Complete nucleotide sequence of bacteriophage T7 DNA and the location of T7 genetic elements. *J Mol Biol* 1983;166:477–535.
23. Epstein RH, Bolle A, Steinberg CM, et al. Physiological studies of conditional lethal mutants of bacteriophage T4D. *Cold Spring Harb Symp Q Biol* 1963;28:375–394.
24. Feiss M, Sippy J, Miller G. Processive action of terminase during sequential packaging of bacteriophage λ chromosomes. *J Mol Biol* 1985;186:759–771.
25. Fitch WM, Leiter JME, Li X, Palese P. Positive Darwinian evolution in human influenza A virus. *Proc Natl Acad Sci USA* 1991; 88:4270–4274.
26. Flemming M, Deumling B, Kemper B. Function of gene 49 of bacteriophage T4. III. Isolation of Holliday structures from very fast-sedimenting DN A. *Virology* 1993;196:910–913.
27. Franklin NC. Altered reading of genetic signals fused to the *N* operon of bacteriophage λ: genetic evidence for modification of polymerase by the protein product of the *N* gene. *J Mol Biol* 1974;89:33–48.
28. Geiduschek EP. Regulation of expression of the late genes of bacteriophage T4. *Annu Rev Genet* 1991;25:437–460.
29. Goliger JA, Roberts JW. Sequences required for antitermination by phage 82 Q protein. *J Mol Biol* 1989;210:461–471.
30. Goyal SM, Gerba CP, Bitton G, eds. *Phage ecology*. New York: John Wiley & Sons, 1987.
31. Haggard-Ljungquist E, Halling C, Calendar R. DNA sequences of the tail fiber genes of bacteriophage P2; evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. *J Bacteriol* 1992;174:1462–1977.
32. Haruna I, Spiegelman S. Specific template requirements of RNA replicases. *Proc Natl Acad Sci USA* 1965;54:579–587.
33. d'Herelle F. Sur un microbe invisible antagoniste des bacilles dysenteriques. *Compt Rend Acad Sci* 1917;165:373–375.
34. Herenden DR, Kassavetes GA, Geiduschek EP. A transcriptional enhancer whose function imposes a requirement that proteins track along DNA. *Science* 1992;256:1298–1303.
35. Hershey AD, Chase M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* 1952;36:31–56.
36. Hershey AD, Rotman R. Linkage among genes controlling inhibition of lysis in a bacterial virus. *Proc Natl Acad Sci USA* 1948;34:89–96.
37. Kacian DL, Mills DR, Kramer FR, Spiegelman S. A replicating RNA molecule suitable for a detailed analysis of extracellular evolution and replication. *Proc Natl Acad Sci USA* 1972;69:3038–3042.
38. Karlin S, Burge C, Campbell AM. Statistical analysis of counts and distributions of restriction sites in DNA sequences. *Nucleic Acids Res* 1992;20:1363–1370.
39. Kashlev M, Nudler E, Goldfarb A, White T, Kutter E. Bacteriophage T4 Alc protein: a transcription termination factor sensing local modification of DNA. *Cell* 1993;75:147–154.
40. Katsura I, Hendrix R. Length determination in bacteriophage lambda tails. *Cell* 1984;39:691–698.
41. Kesteleijn RA, Remaut E, Fiers W, van Duin J. Lysis gene expression of RNA phage MS2 depends on a frameshift during translation of the overlapping coat protein gene. *Nature* 1982;285:35–41.
42. Kitts PA, Nash HA. Homology-dependent interactions in phage site-specific recombination. *Nature* 1987;329:346–348.

42a. Landy A. Dynamic, structural and regulatory aspects of λ site-specific recombination. *Amer Rev Biochem* 1989;58:913–949.

43. Lederberg EM, Lederberg J. Genetic studies of lysogenicity in *Escherichia coli*. *Genetics* 1953;38:51–64.
44. Lindquist BH, Dehö G, Calendar R. Mechanisms of genome propagation and helper exploitation by satellite phage P4. *Microbiol Rev* 193;57:683–702.
45. Lindsey DR, Mullin DA, Walker JR. Characterization of the cryptic lambdoid prophage DLP12 of *Escherichia coli* and overlap of the DLP12 integrase gene with the tRNA gene *argU*. *J Bacteriol* 1989;171: 6197–6205.
46. Little JW. Mechanism of specific LexA cleavage: autodigestion and the role of RecA coprotease. *Biochimie* 1991;73:411–422.
47. Luftig RB, Wood WB, Okinawa R. Bacteriophage T4 head morphogenesis. On the nature of gene 49-defective heads and their role as intermediates. *J Mol Bol* 1971;57:553–573.
48. Luzzati D. Regulation of λ exonuclease synthesis: role of the N gene product and I repressor. *J Mol Biol* 1970;49:515–519.
49. Lwoff A. Lysogeny. *Bacteriol Rev* 1953;17:269–337.
50. Mason SW, Li J, Greenblatt J. Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. *J Biol Chem* 1992;267:19418–19426.
51. Mizuuchi K, Craigie R. Mechanism of bacteriophage Mu transposition. *Annu Rev Genet* 1986;20:385–430.
52. Muller HJ. The relation of recombination to mutational advance. *Mutation Res* 1964;1:2–9.
53. Novick A, Szilard L. Virus strains of identical phenotype but different genotype. *Science* 1951;113:34–35.
54. Oberto J, Weisberg RA, Gottesman ME. Structure and function of the *nun* gene and the immunity region of the lambdoid phage HK022. *J Mol Biol* 1989;207:675–693.
55. Orsine G, Brody EN. Phage T4 DNA codes for two distinct 10 kDa proteins whicb strongly bind to RNA polymerase. *Virology* 1988; 162:397–405.
56. Parma DH, Snyder M, Sobolecsh S, Mowray M, Brody E, Gold L. The Rex system of bacteriophage λ: tolerance and altruistic cell death. *Genes Dev* 1992;6:497–510.
57. Ptashne M. *A genetic switch*. Cambridge, MA: Cell Press and Blackwell; 1992.
58. Riley M, Anilionis A. Conservation and variation of nucleotide sequences within related bacterial genomes: Enterobacteria. *J Bacteriol* 1980;143:366–376.
59. Salas M. Phages with protein attached to their DNA ends. In: Calendar R, ed. *The bacteriophages*. New York: Plenum; 1988:169–192.

60. Schlesinger M. The Feulgen reaction of the bacteriophage substance. *Nature* 1936;138:508–509.
61. Schmieger H, Backhaus H. The origin of DNA in transducing particles in P22-mutants with increased transduction-frequencies (HT-mutants). *Mol Gen Genet* 1973;120:181–190.
62. Scott JR. Sex and the single circle: conjugative transposons. *J Bacteriol* 1992;174:6005–6010.
63. Séchaud J, Streisinger G, Emrich J, et al. Chromosome structure in phage T4. II. Terminal redundancy and heterozygosis. *Proc Natl Acad Sci USA* 1965;54:1333–1339.
64. Son M, Watson RH, Sewer P. The direction and rate of phage T7 DNA packaging *in vitro*. *Virology* 1993;196:282–289.
65. Streisinger G, Emrich J, Stahl MM. Chromosome structure in phage T4. III. Terminal redundancy and length determination. *Proc Natl Acad Sci USA* 1967;57:292–295.
66. Twort FW. An investigation on the nature of ultramicroscopic viruses. *Lancet* 1915;189:1241–1243.
67. Van Etten, JL, Burbank DE, Cuppels PA, Lane LC, Vidaver AK. Semiconservative replication of double-stranded RNA by a virion-associated RNA polymerase. *J Virol* 1980;33:769–783.
68. Wall JD, Harrison PD. Phage P1 mutants with altered transducing abilities for *Escherichia coli*. *Virology* 1974;59:532–544.
69. Watson JD. Origins of concatameric DNA. *Nature* 1972;239:197–201.
70. Weiss RB, Huang WM, Dunn DM. A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene 60. *Cell* 1990;62:117–126.
71. Winter RB, Morrissey L, Gauss P, Gold L, Hsu T, Karam J. Bacteriophage T4 *regA* protein binds to mRNAs and prevents translation initiation. *Proc Natl Acad Sci USA* 1987;84:7822–7826.
72. Yarmolinsky MB, Sternberg N. Bacteriophage P1. In: Calendar R, ed: *The bacteriophages*. New York: Plenum; 1988:291–418.
73. Young R. Bacteriophage lysis: mechanism and regulation. *Microbiol Rev* 1992;56:430–481.
74. Zahler SA, Temperate bacteriophage of *Bacillus subtilis*. In: Calendar R, ed. *The bacteriophages*. New York: Plenum; 1988:559–592.